一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report

Ethereum 上の不正な DeFi トークン対策研究促進のためのデータセット構築

鈴木惟央利† インミン パ パ†† 吉岡 克成†††

† 横浜国立大学大学院環境情報学府
†† 横浜国立大学先端科学高等研究院
††† 横浜国立大学大学院環境情報研究院 / 先端科学高等研究院

あらまし Ethereum ブロックチェーンでは仮想通貨である ETH の他に，スマートコントラクトと呼ばれるプログラムによって暗号資産である DeFi (Decentralized Finance) トークンを実装することができる．DeFi トークンは他の仮想通貨と交換できる仕組みにより価値をもち，投資目的で扱われたり，ゲーム内の通貨として扱われたりする．現在，この DeFi トークンを悪用した詐欺が多発しており，詐欺トークンへの対策はブロックチェーン上の解決すべき問題の 1 つである．そのため，近年ではトークンの安全性を評価するためトークン解析ツールが提供されるようになっている．しかし，ブロックチェーン上のトークンのライフサイクルは早く，338 万ドルを騙し取った Squid Game [1] などの詐欺トークンのプールの情報や Web 上の情報，トークンのメタデータ，安全性の評価値 (安全性スコア) など，過去のデータが記録されていないことが多い．そこで，本研究では，ブロックチェーン上で生成されるトークンを監視し，これらを複数のトークン解析ツールで継続的に検査した結果や，それらのトークンに紐づくメタデータを収集することで，不正トークン検知技術の研究開発の基礎となるデータセットの構築を目指す．
キーワード 暗号通貨, 詐欺, スマートコントラクト, DeFi トークン

Building a Dataset for Accelerating Researches against Fraudulent DeFi Tokens

Iori SUZUKI†, Yin Minn Pa Pa††, and Katsunari YOSHIOKA†††

† Graduate School of Environment and Information Sciences，Yokohama National University
†† Institute of Advanced Sciences，Yokohama National University
††† Graduate School of Environment and Information Sciences/ Institute of Advanced Sciences，Yokohama National University

Abstract In the Ethereum blockchain, in addition to the cryptocurrency ETH, it is possible to implement DeFi (Decentralized Finance) tokens, which are crypto assets, through programs called smart contracts. DeFi tokens have value due to their mechanism of being exchangeable for other cryptocurrencies and are used for investment purposes or as in-game currency. Currently, there is a surge in scams abusing these DeFi tokens, and countermeasures against these fraudulent tokens are one of the problems to be solved on the blockchain. Therefore, in recent years, token analyzers to evaluate the safety of tokens have been emerged. However, the lifecycle of tokens on the blockchain is fast, and often, past data such as pool information for fraudulent tokens like Squid Game [1], which deceived $3.38 million, as well as web information, token metadata, evaluation value (safety score), etc., are not recorded. Therefore, in this study, we aim to build a dataset that will serve as the basis for research and development of fraudulent token detection technology by monitoring tokens generated on the blockchain, continuously inspecting them with multiple token analyzers, and collecting metadata associated with those tokens.
Key words Cryptocurrency, Fraud, Smart Contract, DeFi Token

1. まえがき

近年のブロックチェーンの普及により，仮想通貨の種類が爆発的に増加している．特に Ethereum ブロックチェーンでは仮想通貨である ETH の他に，スマートコントラクトと呼ばれるプログラムによって暗号資産である DeFi (Decentralized Finance) トークンを実装することができる．DeFi トークンは他の仮想通貨と交換できる仕組みにより価値をもち，投資目的で扱われたり，ゲーム内の通貨として扱われたりする．

しかしながら，これらの DeFi トークンを用いた詐欺が多発しており，ブロックチェーン分析企業である Chainalysis が発行した 2023 年のレポートでは，新しいトークンの 4 つに 1 つは詐欺トークンであることが報告されている [2]．また，2024 年には，詐欺により少なくとも 46 億ドルの資金が流出したとされており，その被害は甚大である [3].

詐欺の手法は多様であるが，DeFi トークンを実装しているスマートコントラクト自体にバックドアコードを仕込み，それを用いてスマートコントラクトを改ざんすることでプロジェクトの開発者が募った資金を持ち逃げするハードラグプル，DeFi トークンの開発者がプロジェクトに投資・支援していると見せかけながらもトークンを投棄し，利益を得るソフトラグプル，DeFi トークン開発者グループやプロジェクトのオーナーがトークンのシェアの大半を有している状態で，SNS などにおける情報操作によって資産の価値を不正に引き上げ (パンプ)，価格が高騰した資産を売り抜く (ダンプ) ことで利益を得るパンプアンドダンプ [4] などが知られている．

一方，これらの詐欺への対策として，DeFi トークンおよびスマートコントラクトの安全性を評価する商用のトークン解析ツールである Token Sniffer [5] や DEXTools [6]，Go Plus [7] などが存在する．ブロックチェーン上や Web 上のデータを取得し，それらについて機械学習を行うことで詐欺検知を行う研究 [8] [9] [10] が存在するものの，不正トークン対策に関する学術的な研究は十分に行われているとは言えない．不正トークン対策の研究を行うためには，ブロックチェーン上での DeFi トークンの生成状況や利用状況，作成者，ソースコード，価格情報，公式サイト，関連 SNS アカウント情報といった多様な関連情報を蓄積した研究用データセットを構築し，これらの基礎データに基づき，分析・検知手法を比較・検討することが重要である．また，詐欺トークンの検知手法を検討する上で，既存のトークン解析ツールや検知手法による検知結果も重要な参考情報となる．しかしながら，詐欺の検知に必要である情報は，全てが蓄積されているわけではなく時間と共に失われたり，取得が困難になったりする恐れがある．Etherscan [11] などブロックチェーンに関わる様々な情報を蓄積，分析，公開している，ブロックチェーンエクスプローラと呼ばれるサービスやデータベースも存在するが，それぞれが蓄積するデータは完全ではない．

そこで本研究では, ブロックチェーン上で生成されるトークンを監視し, これらの関連情報を収集すると共に, 複数のトークン解析ツールの検査結果を継続的に収集, 蓄積することで, 不正トークン検知技術の研究開発の基礎となるデータセットの構築を目指す.

構築した監視システムを用いて 2024 年 5 月 13 日から 2024 年 6 月 17 日まで観測を実施した結果, 32,316 個のスマートコントラクトが得られた．そのうち，ソースコードの分析により 11,920 個をトークンと判定し，2 種類のトークン解析ツールである Token Sniffer と DEXTools が出力する安全性スコアと検査結果を継続的に取得・蓄積し，データセットを構築した．Token Sniffer は 11,920 個のうち，10,869 個に対してスコアを取得でき，約 39% にあたる 4,213 個のトークンを不正トークンと判定し，約 33% にあたる 3,602 個のトークンにエクスプロイトが含まれると判定した．一方，DEXTools は，11,454 個に対してスコアを取得でき，約 1.6% にあたる 188 個のトークンを詐欺の可能性があるトークンであると判定した．

これらのデータセットの有用性を確かめるため，簡易的な分析として安全性スコアの時間的変化や 2 つのトークン解析ツールの安全性スコアの関係性を調査した．その結果, トークン生成から 50 時間程度で安全性スコアはある程度安定するものの，Token Sniffer については一部のトークンについて生成から 200 時間以上経過した後にも大きなスコア変更がある場合が確認できた．また 2 つのツールの判定には大きな差があることがわかった．

2. 研究背景

2.1 用語解説

2.1.1 ブロックチェーン

ブロックチェーンは, 取引記録であるトランザクションがまとめられたブロックが一本の鎖状に繋がっている分散型のデータベースである. 各ブロックには前のブロックのハッシュ値を含んでおり, P2P ネットワークを利用しているため, 改竄や攻撃に対して強いという特徴を持つ. ブロックチェーン上のデータは暗号化されており, ブロックチェーンに接続した複数のノードに分散して保存されているため, セキュリティや信頼性が高い. ブロックチェーン上の情報は誰でも閲覧可能であり, ブロックチェーンエクスプローラなどを用いて, 情報の確認が可能である.

ブロックチェーン技術は Bitcoin [12] や Ethereum [13] などに用いられているが, 仮想通貨だけでなく, 投票システムや製品管理, 証明書の発行にも応用されている. また, Ethereum では ETH と呼ばれる資産とトークンがやり取りされており, ブロックチェーン上で実行されるプログラムであるスマートコントラクト (以下コントラクト) を実装可能である.

本研究では Ethereum ブロックチェーンを対象とする.

2.1.2 ブロックチェーンエクスプローラ

ブロックチェーンエクスプローラとは特定のブロックチェーンネットワーク上でのトランザクションやブロックなどの情報を継続的に取得, 蓄積し, 提供するウェブベースのツールを指す．Ethereum においては, Etherscan [11] が代表的なブロックチェーンエクスプローラである.

このEtherscanでは，ソースコードの透明性を確保するために，コントラクトの開発者がソースコードを公開している場合がある．これをコントラクトの検証 (Verify) と呼ぶ．コントラクトの検証は，Etherscanに送信したソースコードのバイトコードとブロックチェーン上のコントラクトのバイトコードが一致しているかどうかを確認することにより行われる．

また検証には通常開発者しか知り得ないコントラクトのソースコードとコンストラクタ引数を必要とするため，第三者が勝手にコントラクトの検証を行うことは原則として困難である．

2.1.3 スマートコントラクト

スマートコントラクトは，デジタルな方法によりあらゆる資産が動的に処理されるような契約や，取引における契約および執行を自動で行う仕組みやプログラムのことを指す. これにより，仲介者なしでの取引が可能になる [14]. 一般的に，ブロックチェーン上で用いられているものを指し，賭け事や不動産の取引，ゲーム，投票システムなどにも使用されている. スマートコントラクトは，Ethereumなどのブロックチェーン上で用いられており，所定の条件が満たされるとブロックチェーンにトランザクションを自動的に書き込む.

スマートコントラクトを用いて，独自の暗号通貨やデジタル資産を意味するトークンを実装することができる．また，スマートコントラクトはアドレスと呼ばれる一意の文字列をもつ．

2.1.4 トークン

ブロックチェーン上で発行されるデジタル資産を指す．これらのトークンは，スマートコントラクトによって管理され，Ethereumネットワーク上で取引できる．Ethereumの機能やプロセスの規格はEIPs (Ethereum Improvement Proposals) [15] で提案され，ERC (Ethereum Request for Comments) とよばれる技術文書で概説されており，ERCにはトークンに関する規格も含まれる．Ethereum上のトークンには様々な規格があり，それぞれ異なる機能や性質をもつ．

ERC-20 [16]: 最も一般的に使用されるトークンの規格で，トークンの発行や転送，残高の確認など基本的な機能を定義している．この規格に基づいたトークンはEthereumネットワーク上の異なるウォレットや取引所と互換性があり，代替可能なもの (Fungible Token) である．

ERC-721 [17]: 非代替可能なトークン（Non-Fungible Token) NFTの標準規格であり，各トークンは独自の識別子を持つため，異なる価値や属性を表現できる．すなわち，この規格のトークンの複製や偽造は困難である．この規格は，デジタルアートや仮想世界のアセットなどで主に利用されている．

2.1.5 DeFiトークン

DeFiトークンはERC-20トークンの一種である．分散型金融（DeFi）プラットフォームやアプリケーション内で使用される暗号通貨で，イーサリアムなどのブロックチェーン上で動作し，金融サービスの運用と発展を支えている．代表的な例として，分散型取引所 Uniswap [18] のUNIトークンがある．

2.1.6 流動性プール

流動性プールはトークンの取引に必要な流動性を提供するための仕組みであり，分散型取引所DEX (Decentralized Exchange) に預けられたトークンの集合を指す．Ethereumにおける代表的なDEXとしてUniswap [18] がある．流動性プロバイダLP (Liquidity Provider) は同じ価値をもつ基軸通貨とトークンを流動性プールに預け，流動性を提供する．この流動性プールに基軸通貨を手数料とともに預け入れることでトークンを得ることができる．トークンを売却するときは，同様にトークンを手数料とともに流動性プールに預け入れることで，同価値の基軸通貨を得ることができる．LPは提供した流動性に応じて，その手数料の一部を得ることができる．

また，流動性にはロックをかけることができ，LPが預けた基軸通貨とトークンのセットを一定期間引き出せないようにする流動性ロックと呼ばれる仕組みが提供されている．これにより，トークンの価値が暴落することを防ぐことができる．この仕組みはLPロックと呼ばれている．多くの分散型金融 (DeFi) プラットフォームでは，売り手と買い手による従来の市場の代わりにAMM (Automated Market Makers) が使用されており，流動性プールを使用してデジタル資産を自動的に取引できる [19].

2.1.7 ラグプル

ラグプルは，誇大広告によってトークンの価格を吊り上げた後，トークンの開発者が資金を持ち逃げする詐欺を指す [20]. 出口詐欺と呼ばれる場合もある．ラグプルにはハードラグプルとソフトラグプルの2種類が存在する．

ハードラグプル: ハードラグプルはスマートコントラクト自体にバックドアを仕込み，それを用いてスマートコントラクトを改ざんすることでプロジェクトの開発者が募った資金を持ち逃げする手法を指す．不正度が高く，犯罪行為である場合がほとんどである．Squid Gameトークンの詐欺 [1] ではこのハードラグプルが発生し，開発者は約338万ドルを持ち逃げした．

ソフトラグプル: 一方，ソフトラグプルは，トークンの開発者がトークンを素早く投棄する手法を指す．トークンの価格を情報操作などによって意図的に吊り上げた後，流動性プール内の開発者自身が保有するトークンを売り抜くことによって利益を得る手法であるパンプアンドダンプはこれに該当する．また，LPが事前の宣言なしに流動性プールから基軸通貨とトークンのセットを引き出し，それを売却することで利益を得る手法もソフトラグプルの内の1つである．

ラグプルでは，開発者はSNSやWebページを主体として投資家に投資を促している [21]．トークンの開発者は，トークンと関連するコミュニティにアクセスするためのSNSリンクをスマートコントラクトのソースコードに記載することで，投資家をX(Twitter) やTelegramなどのコミュニティに誘導する．

2.1.8 不正トークン検知ツール

Token Sniffer [5]: Token Sniffer は，ブロックチェーン上の新しいトークンをスキャンし，その信頼性やセキュリティリスクを評価するツールである．トークンのリスクを 0 以上 100 以下の整数値の安全性スコアで表現し，安全であると推測されるトークンほどスコアが高く表示される．Token Sniffer はコントラクトがデプロイされると取引の分析やコントラクトのソースコード，流動性などを自動的に分析し，安全性スコアを算出する．トークンを実装するコントラクトのアドレスを Web ページで入力することで，トークンの評価を取得することができる．また，提供されている API を使ってトークンの監査結果を取得することもできる．

DEXTools [6]: DEXTools は，分散型取引所（DEX）におけるトレーディングおよび市場分析ツールを提供するサービスである．市場分析や取引記録の可視化などの機能以外に，トークンの信頼性を表現した DEXTScore [22] と呼ばれるアルゴリズムを提供している．DEXTScore はトークンのリスクを 0 以上 99 以下の整数値の安全性スコアで表現し，安全であると推測されるトークンほどスコアが高く表示される．DEXTScore で測定されるパラメータには，ソーシャルメディアからの情報，トークンペアの流動性，トランザクション数，トークン保有者数などがある．

本研究では，トークンのリスクが安全性スコアとして提供されている Token Sniffer と DEXTools を用いる．

2.2 関 連 研 究

2.2.1 詐欺トークン検知

Ethereum 上の詐欺トークンを検知するための研究は以前から行われており，様々なアプローチで詐欺トークンを検知している．これらのアプローチは基本的にはトランザクションなどのブロックチェーン上の情報を分析するものが多い．

Jiajing Wu らの研究 [8] では，トランザクションを利用してフィッシングアドレスを識別している．各アドレスの特徴を抽出するために取引額とタイムスタンプを考慮したネットワーク埋め込みアルゴリズム（trans2vec）を提案しており，埋め込みベクトルを入力として，one-class SVM を用いてフィッシングアドレスと非フィッシングアドレスを分類している．

他にも詐欺トークンの検知に関連する研究として，Ifeyinwa Jacinta Onu らによる機械学習を用いたポンジスキームの検出に関する研究 [9] などがある．

また，Kentaro Asaba による研究 [10] では，仮想通貨のプロジェクトの内容について説明するホワイトペーパーを入力として，機械学習により詐欺トークンを予測している．

2.2.2 不正トークンのデータセット

Pengcheng Xia らの研究 [23] では，分散型取引所である Uniswap 上の詐欺トークンに関して大規模なデータセットを構築し，それらについて分析を行なっている．同様に，Suparat Srifa らの研究 [24] では，オンチェーンのデータを含む，Uniswap 上のトークンの情報を取得し，機械学習を行うことによって，ラグプル発生の実態を明らかにし，予測を行なった．これらの研究は Uniswap のみを対象としているが，本研究では Ethereum 上の全てのトークンを監視しているため，Uniswap 以外の DEX にも対応する．

また Sebastian Luzian による研究では [25] では，Web 上で公開されている様々なデータソースから不正だとラベルが付けられたコントラクトアドレスを収集し，ブロックチェーン上のデータを用いて不正トークンの取引フローを明らかにしている．しかしながら，Web 上のデータソースには限りがあり，頻繁に更新されるものでもないため，網羅的に不正トークンについて調査を行うことはできていない．本研究では，Ethereum ブロックチェーンを常に監視しており，網羅的な調査が可能である．

3. データセット作成手法

3.1 概　　要

本研究では，図 1 のような流れに沿って，Ethereum メインブロックチェーン上で新しくデプロイされたコントラクトをモニタリングし，関連するデータを収集することでトークン解析ツールの検証およびデータセットの構築を行う．まず Web3 ライブラリ [26] を用いて Ethereum ノードに接続し，新規にデプロイされたコントラクトのアドレスを収集する．収集したアドレスに対し，ブロックチェーンエクスプローラである Etherscan でそれらのソースコードを開発者が公開しているかをチェックし，ソースコードを収集する．

次に，そのソースコードを解析しトークンの実装に必要な関数が実装されているかどうかを確認し，収集したコントラクトがトークンを実装しているものなのかどうかを判別する．トークンであると確認できたものに関して，Token Sniffer と DEXTools の API を用いて一定間隔ごとに安全性スコアおよびメタデータを取得する．取得したデータから詐欺であったり悪意があったりすると考えられるものに関してラベル付けし，データセットを構築する．メタデータにはトークンの名前やプールに関する情報などが含まれる．

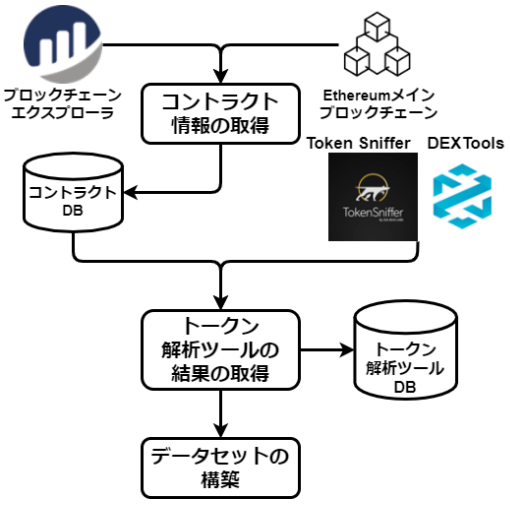

図 1　Overview

3.2 データの収集

3.2.1 コントラクトの収集

本研究のデータ収集期間は 2024 年 5 月 13 日から 2024 年 6 月, 17 日である．

Web3 ライブラリを用いて，Cloudflare から提供されている Ethereum ノード [28] に接続し，各ブロックのトランザクションをチェックする．トランザクションの宛先アドレスが指定されていないものはコントラクトデプロイのためのトランザクションであるため，該当するトランザクションから新規にデプロイされたコントラクトのアドレスを保存する．また，ブロックチェーンエクスプローラである Etherscan の Etherscan API [27] を用いることで，ソースコードやコントラクトが作成された日時，コントラクト名などを取得する．

ソースコードを分析し、totalSupply や allowance などの必須関数が実装されているかどうかを確認することで，トークンを実装しているコントラクトかどうか判別し，ラベルをつける．

一つのトークンに対し，開発者からソースコードが Etherscan 上に公開されたかどうかをを 12 時間ごとに 20 回まで確認する．また，公開が確認できた日時も保存する．

3.2.2 スコアの収集

収集したトークンに関して，トークン解析ツールの API を用いてスコアを 1 時間ごとに取得する．本研究では Token Sniffer API と DEXTools API を用いる．特に，Token Sniffer は Enterprise プラン，DEXTools は Standard プランを用いる．それぞれ制限があり，Token Sniffer は 1 日 5,000 トークンを追跡可能で，DEXTools は月に 1,000,000 回リクエスト (最大 2 リクエスト/秒) が可能である．

Token Sniffer API では，Get Token を用い，得られるデータを全て保存する．得られるデータはトークンのスコア以外にトークンの名称や作成日時，売買手数料などの基本情報, エクスプロイト (詐欺トークンにみられる特徴), トップホルダ (トークンの上位保有者の情報), プール (トークンの LP (Liquidity Pool) に関する情報), プールご毎トップホルダ (流動性プール毎のトークンの上位保有者の情報), 流動性ロック情報 (ユーザーが提供した資金がどれだけの期間または条件でロックされているかの情報), ソースコードの類似度 (他のトークンとの類似度), 各監査の結果 (監査のテスト内容と実行結果) である．

また，DEXTools API ではアドレスや名前などのトークンの詳細，詐欺かどうかなどの監査情報，トランザクション数などの追加の情報，流動性ロック情報，トークンのスコアなどを取得する．

3.3 データセットの構築

収集したデータにラベル付けしてデータセットを構築する．データセットはコントラクトの名前, トークンの名前, ソースコード, コントラクトの作成日時, コントラクトアドレス, クリエイタアドレス (コントラクトの作成者を特定するための一意の文字列), トークンの価格に関する情報 (トークンの総供給量や価格変化，購入者などの情報), プールに関する情報 (トークンがどの DEX で取り扱われているかなどの情報)，安全性スコアで構築した．

4. 結　　果

4.1 コントラクトの収集

コントラクトのモニタリングは 2024 年 5 月 13 日 11:20 (UTC) から開始し，表 1 のような結果が得られた．2024 年 6 月 17 日 00:00 (UTC) 時点で 31,955 個であり，そのうちトークンであると判定されたものは 11,725 個 であった．ただし，単に必須関数が実装されているか確認するという簡素な方法でトークンかどうかを判定しているため，トークンではないコントラクトが含まれている可能性がある．また，プログラムのバグによる影響でコントラクトのデプロイから情報を取得するまでの時間が長いものがあり，それらのコントラクトに関してはトークンのスコアデータが一部欠損している．

4.2 トークンのスコア追跡

Token Sniffer のスコアの追跡は 2024 年 5 月 13 日 11:28 (UTC) から開始し, 2024 年 6 月 17 日 00:00 (UTC) 時点で 10,739 個のトークンについて追跡を行なっており, 1,022,885 個のレコードが得られた．スコア追跡のためのプログラムのバグの影響で，一部データの欠損期間があり，2024 年 5 月 21 日 01:09(UTC) から 2024 年 5 月 25 日 21:19(UTC) はデータが欠損している．また，DEXTools のスコアの追跡は 2024 年 5 月 28 日 21:00 (UTC) から開始し，2024 年 6 月 17 日 00:00 (UTC) 時点で 12,979 個 のトークンについて追跡を行なっており．247,708 個 のレコードが得られている．こちらも API のリクエスト制限とプログラムのバグにより，追跡を開始してから 2024 年 6 月 10 日 02:50 (UTC) まではデータが欠損しており，完全ではない．

基礎的な統計量は表 2 のようになった．

表 1　コントラクトに関する統計

収集期間 (UTC)	2024/05/13 - 2024/06/17
欠損期間 (UTC)	2024/05/15 - 2024/05/28 2024/06/04 - 2024/06/10
コントラクト総数	32,316
トークン総数	11,920
平均コントラクトデプロイ数 (/日)	864
平均トークンデプロイ数 (/日)	313
平均クリエイタ (/日)	587

5. 分析の一例

5.1 用 語 定 義

デッドトークン: 本研究では，トークン解析ツールにおけるスコアを追跡するが，安全性スコアが継続的に最低値を示すものはトークンとしての役割を終えていると判定し，デッドトークンと呼称する．デッドトークンとして判定する際，一週間の範囲内で判定した時刻以降のスコアが全て最低値を取っているかどうかを確認する．最低値ではないスコアがある場合はデッドトークンのラベルはつけない．

表 2　トークンに関する統計

項目	Token Sniffer	DEXTools
追跡したトークン数	10,869	11,454
プールが存在するトークン数	9,292	4,272
不正なトークンの ラベルをつけられたトークン数	4,213	-
エクスプロイトがあったと ラベル付けされたトークン数	3,602	-
詐欺である可能性がある トークン数	-	180

ライフタイム: トークンを実装するコントラクトがデプロイされてからデッドトークンと判定されるまでの時間をトークンのライフタイムと呼称する．ライフタイムは時間単位で測定し，コントラクトがデプロイされて以来スコアが 0 になったことがないトークンは 1 週間以上 (168 時間以上) の枠組みに入れる．

5.2　スコアの時間変化

Token Sniffer のスコアの時間変化を 4 つのパターンに分類し，図 2 が得られた．図 2 は，各分類からランダムに 10 個を選んで表示している．

スコア一定型トークン (黄): コントラクトがデプロイされてからずっとスコアが一定か，ほぼ変化していないトークン

スコア変動型トークン (青): スコアが時間によって激しく変動しているトークン

スコア下降型トークン (赤): スコアが非ゼロ値からある時点で 0 になったトークン

スコア上昇型トークン (緑): スコアが 0 から非ゼロ値に上昇したトークン

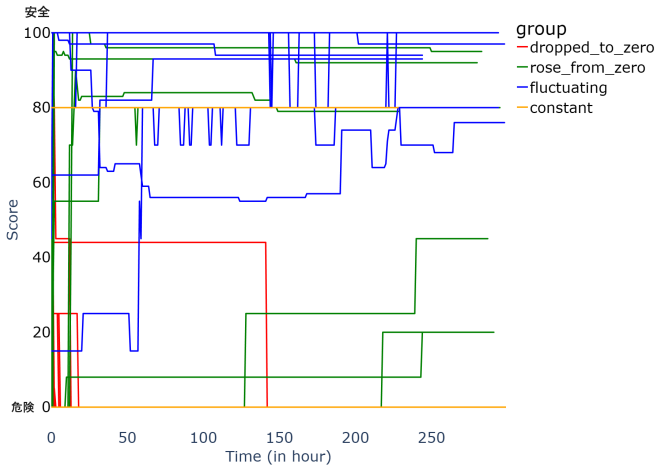

図 2　TokenSniffer のスコアの時間変化

DEXTools のスコアの時間変化も同様に 4 つに分類したところ，図 3 のような結果が得られ，各分類におけるトークン数は表 3 のようになった．

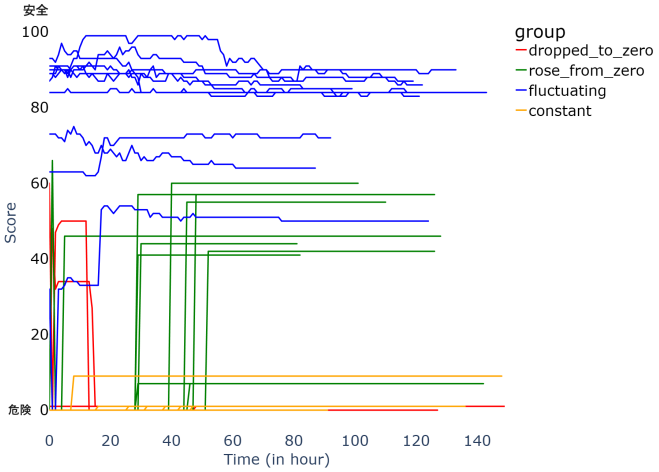

図 3　DEXTScore の時間変化

表 3　スコアの分類

項目	Token Sniffer	DEXTools
スコア一定型トークン	1,787	1,046
スコア変動型トークン	115	273
スコア下降型トークン	211	1,141
スコア上昇型トークン	262	754

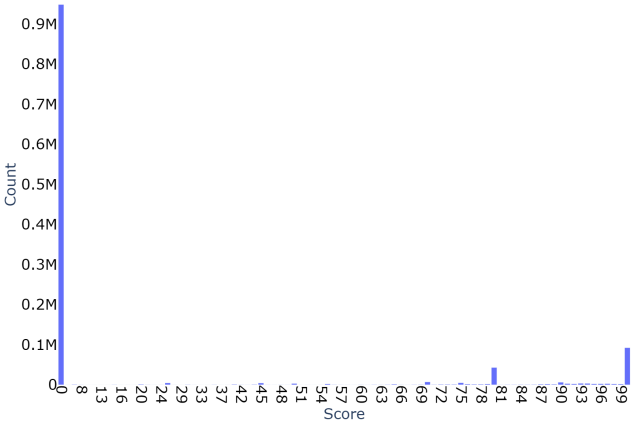

図 4　Token Sniffer の平均スコアの分布

5.3　平均スコアの分布および相関

平均スコアの分布について，Token Sniffer では図 4，DEXTools では図 5 が得られた．両サービスでスコアを取得できたトークンは 711 個であり，2 つのツールのスコアの相関は図 6 のようになった．

5.4　ライフタイム

Token Sniffer と DEXTools で調べたトークンのライフタイムはそれぞれ図 7 と図 8 のようになった．Token Sniffer は 2 時間以内のものが最も多く，その個数は全体の 125 個に対して 25 個であった．一方，DEXTools は 168 時間 (7 日) 以上のものが最も多く，その個数は全体の 326 個に対して 64 個であった．

5.5　全体的な傾向

トークンのスコアは初期段階で 0 になるものが多く，ソフ

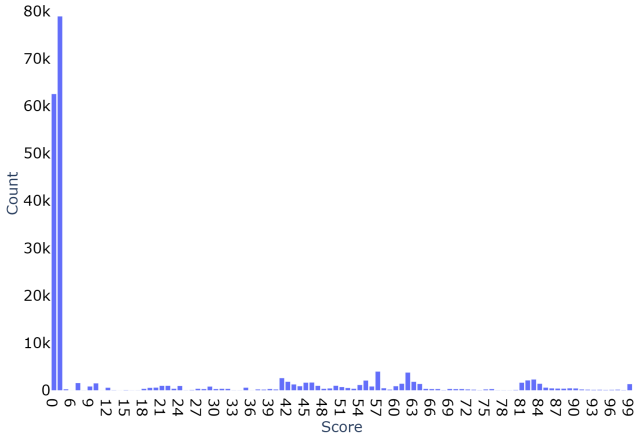

図 5　DEXTools の平均スコアの分布

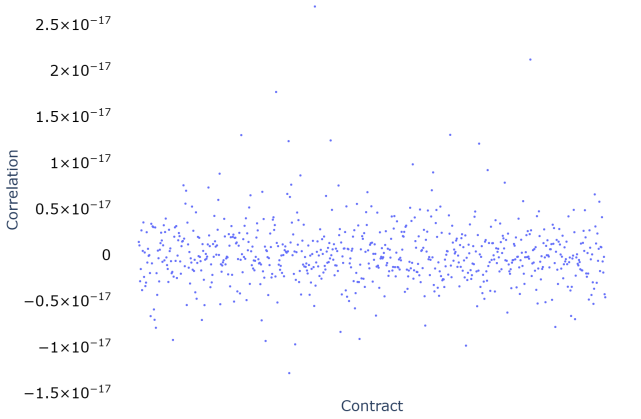

図 6　スコアの相関

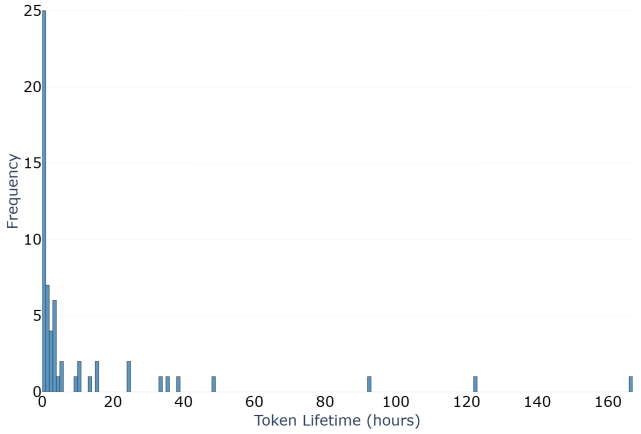

図 7　Token Sniffer のライフタイム

トラグプルが発生する可能性のあるトークンもあることを考慮すると，正当なトークンはわずかしかないことがわかった．また 2 つのツールの間で安全性スコアの相関はほぼなく，各ツールは独自の判定を行っていると言える．

5.6　不正トークン検知ツールの検知結果の特徴

図 4 と図 5 から，両サービスで平均スコアの分布も大きく異なることがわかる．Token Sniffer におけるスコアについては，スコアがトークンのコントラクトデプロイから 100 時間以上経った後に 0 から上昇していたり，150 時間程度経過後に 0 に下降していたりするなど，ある時点での急激な変化が見られる．また，0 から上昇しているトークンの内，10〜20 時間以内に上昇したものは，高いレベルのスコアで安定しており，スコアが常に変動しているトークンは，50 時間程度経過すると安定する傾向にある．DEXTools の結果に関しても，50 時間を超えると概ね評価が大きく変わることはなく，デプロイ時点で高いレベルのスコアをもつトークンは安定してスコアが高い．一方，Token Sniffer については，生成から 200 時間以上経過したトークンについてもスコアが急激に変化する場合もある．以上より，両ツールの結果は，時間経過とともにある程度安定するものの，短期間で安全性を確実に判定できるものではないことが分かった．

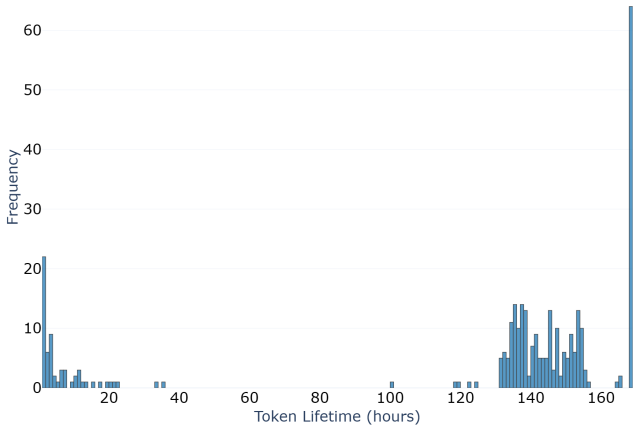

図 8　DEXTools のライフタイム

6.　考　　察

6.1　API 制 限

本研究では，約一ヶ月という短い期間でデータ収集を行ったが，ソフトラグプルといった長期間にわたる詐欺行為を検知するためにはより長期間スコアの追跡を行う必要がある．

しかしながら，現時点では API 制限のため，日々増加するトークンを全て長期間にわたって追跡することは困難である．Token Sniffer では 5,000 個より多くのトークンを監視することができず，DEXTools は月間リクエスト数および秒間リクエスト数の制限があり，同期的処理が必要であった．監視対象のトークンは次第に増えていくため，1 週間以上デッドトークンであるものや，どちらのサービス上でも監査情報がないものを除外することなどを検討すべきである．

6.2　データセットに追加すべき情報

ライフタイムが短いトークンが多いことを考えると，早期に消えてしまうデータを優先的にデータセットに含めるべきである．例えば Web ページやホワイトペーパーは，投棄されたトークンの場合は維持が不要であるため，優先すべきデータである．そのためにもソースコードに記載された URL などから随時 Web ページのアーカイブを取る必要がある．これにより，仮想通貨のプロジェクトの計画が記載されたホワイトペーパーが取得でき，Kentaro Asaba の研究 [10] のようなホワイトペーパーを用いた詐欺の検知に貢献できる可能性がある．

SNS コミュニティも同様で，消えてしまう可能性が高い上，トークンの投棄と SNS の活動は連動している可能性があるため，SNS コミュニティのメッセージなども可能であれば含めるべきである．

以上より，新たにデータセットに含めるべきデータとして Web データと各種 SNS データを提案し，今後収集を行う．

6.3 取得した安全性スコアについて

本研究で取得した安全性スコアは基本的にデータセットに含めるが，それを広く公開したり，共有する場合にはツールの利用規約上注意が必要である．

7. 結　　論

構築したデータセットを用いて異常検知モデルや分類モデルを構築することで，将来的に不正トークンの高精度な予測が可能になる可能性がある．

本研究は，トークンのセキュリティ評価における複数の情報源の信頼性を示すとともに，詐欺検知のための新しいデータセットの有用性を強調している．今後の研究では，これらのデータを活用したモデルの精度向上や，新たな詐欺検知手法の開発が期待される．

8. 今後の展望・計画

詐欺検知のための礎となるデータセットを構築するには，現状では短期間のデータしかないため，不十分である．

また，他のサービスについても追跡すべきであるがスコアのような共通の指標がないため，手法を検討する必要がある．

以上より，今後の計画としては，次の点に焦点を当てる．

（1） 追跡対象トークンの選定：長期間にわたって追跡するトークンを選定し，除外する条件を明確にする．これにより，データの一貫性と信頼性を確保する．

（2） 追加データの取得：他のトークン解析ツールからのデータも含め，スコア以外の有用なインデックスを検討する．特に，スコアが提供されていないサービスに対しては，独自の評価基準を設ける．

（3） データセットの拡充：Web データや SNS 情報，ホワイトペーパーなど，ブロックチェーン上にない情報を含めた包括的なデータセットを構築する．これにより，より高精度な詐欺検知モデルの開発が可能となる．

謝辞　本研究の一部は JSPS 科研費 22H03588 の助成を受けて行われた．

文　　献

[1] BBC News. ”Squid Game crypto token collapses in apparent scam”. 2021. https://www.bbc.com/news/business-59129466, (Ref. 2024 年 8 月 6 日).

[2] Chainalysis, ”The 2023 Crypto Crime Report (日本語版)”, p.106, Chainalysis, 2023-02. (Ref. 2024 年 8 月 6 日).

[3] Chainalysis, ”The 2024 Crypto Crime Report (日本語版)”, p.104, Chainalysis, 2024-02. (Ref. 2024 年 8 月 6 日).

[4] FoolProofMe. ”Major Crypto Scams Explained: ’Pump and Dump’ vs. ’Rug Pull’”. https://www.foolproofme.org/articles/864-major-crypto-scams-explained-pump-and-dump-vs-rug-pull, (Ref. 2024 年 8 月 6 日).

[5] Solidus Labs. ”Token Sniffer”. Token Sniffer. https://tokensniffer.com/, (Ref. 2024 年 8 月 6 日).

[6] DEXTools.io. ”DEXTools - Join the revolution of DeFi”. 2024. https://info.dextools.io/, (Ref. 2024 年 8 月 6 日).

[7] GoPlus Labs. ”GoPlus”. 2024. https://gopluslabs.io/, (Ref. 2024 年 8 月 6 日).

[8] Jiajing Wu, Qi Yuan, Dan Lin, Wei You, Weili Chen, Chuan Chen, Zibin Zheng. ”Who Are the Phishers? Phishing Scam Detection on Ethereum via Network Embedding”. IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 52, no. 2, pp. 1156-1166, Feb. 2022.

[9] Ifeyinwa Jacinta Onu, Abiodun Esther Omolara, Moatsum Alawida, Oludare Isaac Abiodun, Abdulatif Alabdultif. ”Detection of Ponzi scheme on Ethereum using machine learning algorithms”. Sci Rep 13, 18403 (2023). https://doi.org/10.1038/s41598-023-45275-0.

[10] Kentaro Asaba. ”Scam Cryptocurrency Detections using Machine Learning Techniques”. 2019 年度 人工知能学会全国大会 (第 33 回).

[11] Ethereum. ”Etherscan”. https://etherscan.io/, (Ref. 2024 年 8 月 6 日).

[12] Satoshi Nakamoto. ”Bitcoin: A Peer-to-Peer Electronic Cash System”. 2008. https://bitcoin.org/bitcoin.pdf, (Ref. 2024 年 8 月 6 日).

[13] Vitalik Buterin. ”Ethereum White Paper: A Next Generation Smart Contract & Decentralized Application Platform”. 2023-08-16. https://ethereum.org/en/whitepaper/, (Ref. 2024 年 8 月 6 日).

[14] Ethereum. ”INTRODUCTION TO SMART CONTRACTS”. https://ethereum.org/en/developers/docs/smart-contracts/, (Ref. 2024 年 8 月 6 日).

[15] Ethereum Foundation. ”Ethereum Improvement Proposals (EIPs) – ERC”. 2023. https://eips.ethereum.org/erc, (Ref. 2024 年 8 月 6 日).

[16] Ethereum Foundation. ”Ethereum Improvement Proposals (EIPs) – ERC-20: Token Standard”. 2015-11-19. https://eips.ethereum.org/EIPS/eip-20, (Ref. 2024 年 8 月 6 日).

[17] Ethereum Foundation. ”Ethereum Improvement Proposals (EIPs) – ERC-721: Non-Fungible Token Standard”. 2018-01-24. https://eips.ethereum.org/EIPS/eip-721, (Ref. 2024 年 8 月 6 日).

[18] Uniswap Labs. ”Uniswap Protocol”. https://uniswap.org/, (Ref. 2024 年 8 月 6 日).

[19] Cryptopedia. ”What Are Liquidity Pools?”. 2023-11-17. https://www.gemini.com/cryptopedia/what-is-a-liquidity-pool-crypto-market-liquidity, (Ref. 2024 年 8 月 6 日).

[20] Frank Corva. ”What is a crypto rug pull?”. Finder. 2022-11-15. https://www.finder.com/cryptocurrency/what-is-a-crypto-rug-pull, (Ref. 2024 年 8 月 6 日).

[21] biswap BLOG. ”What’s Rug Pull & How to Keep Safe from It?”. 2024-02-14. https://blog.biswap.org/article/what-is-rug-pull, (Ref. 2024 年 8 月 6 日).

[22] DEXTools.io. ”Crypto Glossary - DEXTScore”. 2024. https://info.dextools.io/crypto-glossary/dextscore/, (Ref. 2024 年 8 月 6 日)

[23] Pengcheng Xia, Haoyu Wang, Bingyu Gao, Weihang Su, Zhou Yu, Xiapu Luo, Chao Zhang, Xusheng Xiao, Guoai Xu. ”Trade or Trick?: Detecting and Characterizing Scam Tokens on Uniswap Decentralized Exchange”. Proceedings of the ACM on Measurement and Analysis of Computing Systems. Volume 5, Issue 3, Article No.: 39, pp. 1 – 26.

[24] Suparat Srifa, Yury Yanovich, Ahmad Salehi S., Robert Vasilyev, Tharuka Rupasinghe, Vladislav Amelin. ”Scam Token Classification for Decentralized Exchange Using Transaction Data”. 2023-09-25. https://ssrn.com/abstract=4582918, (Ref. 2024 年 8 月 6 日).

[25] Sebastian Luzian. ”A Systematic Investigation of Illicit Money Flows in the DeFi Ecosystem”. 2022-11-10.

[26] Ethereum. ”web3.py”. GitHub repository. https://github.com/ethereum/web3.py, (Ref. 2024 年 8 月 6 日).

[27] Ethereum. ”Etherscan APIs”. https://etherscan.io/apis, (Ref. 2024 年 8 月 6 日).

[28] Cloudflare. ”Cloudflare Web3 Docs - Ethereum Gateway”. https://developers.cloudflare.com/distributed-web/ethereum-gateway/, (Ref. 2024 年 8 月 6 日).